Strain Engineering of the Band Gap of HgTe Quantum Wells Using Superlattice Virtual Substrates

Philipp Leubner,* Lukas Lunczer, Christoph Brüne, Hartmut Buhmann, and Laurens W. Molenkamp
Experimentelle Physik III, Physikalisches Institut, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany
(Received 9 May 2016; published 19 August 2016; corrected 11 August 2017)

The HgTe quantum well (QW) is a well-characterized two-dimensional topological insulator (2D TI). Its band gap is relatively small (typically on the order of 10 meV), which restricts the observation of purely topological conductance to low temperatures. Here, we utilize the strain dependence of the band structure of HgTe QWs to address this limitation. We use CdTe-$Cd_{0.5}Zn_{0.5}Te$ strained-layer superlattices on GaAs as virtual substrates with adjustable lattice constant to control the strain of the QW. We present magneto-transport measurements, which demonstrate a transition from a semimetallic to a 2D-TI regime in wide QWs, when the strain is changed from tensile to compressive. Most notably, we demonstrate a much enhanced energy gap of 55 meV in heavily compressively strained QWs. This value exceeds the highest possible gap on common II-VI substrates by a factor of 2–3, and extends the regime where the topological conductance prevails to much higher temperatures.

DOI: 10.1103/PhysRevLett.117.086403

The transport properties of molecular-beam epitaxially (MBE) grown HgTe quantum wells (QWs) embedded in $Cd_{0.7}Hg_{0.3}Te$ barriers have attracted considerable attention due to the discovery of the quantum-spin-Hall (QSH) effect in these structures [1–3]. The QSH effect is the landmark property of a two-dimensional topological insulator (2D TI) and is characterized by the presence of a pair of one-dimensional, counterpropagating ("helical") channels along the edges of the mesa, giving rise to a quantized longitudinal conductance $G_{\mathrm{QSH}} = e^2h^{-1}$ [1]. A prerequisite for the formation of edge channels is a—topologically nontrivial—inverted band structure, as is present in HgTe QWs when the thickness d_{QW} exceeds $d_c = 6.3$ nm [1]. Inverted HgTe QWs have a relatively small band gap E_G (typically lower than 15 meV), which can make it difficult to gate homogeneously into the gap over the whole mesa, and also prevents applications at elevated temperatures. Here, we present a way to increase E_G well above the thermal energy at room temperature ($k_BT = 25$ meV). This is achieved by applying compressive strain to HgTe QWs through coherent growth on virtual substrates with a freely tunable lattice constant.

The crucial influence of strain on the band structure of HgTe has been demonstrated previously for bulk layers (layer thickness $d > 40$ nm): epitaxy of HgTe on CdTe substrates exerts tensile strain ($\varepsilon = -0.3\%$), which causes a gap opening of the Γ_8 doublet, transforming the bulk semimetal into a three-dimensional topological insulator [4,5]. However, these previous experiments used commercially available MBE quality substrates, limiting the options to $Cd_{0.96}Zn_{0.04}Te$ [1–3] and CdTe [4–6]. In both cases, the lattice constant of the substrate material is larger than that of HgTe, resulting in a tensile strain in the epilayers. Under such conditions, the largest gaps that can be obtained in inverted QWs are $E_G = 17$ and 25 meV for wells grown on CdTe and $Cd_{0.96}Zn_{0.04}Te$, respectively [7].

The present work reports on a major progress in this situation. We use CdTe-$Cd_{0.5}Zn_{0.5}Te(001)$ strained-layer superlattices (SLS) as virtual substrate material for HgTe-based epilayers. These superlattices (grown on a GaAs substrate) provide a straightforward control of the effective lattice constant of the system and, thus, the strain in the subsequently grown HgTe layers. The use of SLS rather than (Cd,Zn)Te solid solutions for lattice constant control is necessary because solid solutions suffer from poor crystal quality due to phase separation effects [8,9]. We fabricate both tensile ($\varepsilon < 0$) and compressively ($\varepsilon > 0$) strained QWs using coherent epitaxy of (Zn,Cd,Hg)Te-HgTe-(Zn,Cd,Hg)Te heterostructures on virtual substrates. High-resolution x-ray diffraction (HRXRD) is used to analyze the SLS crystal structure and to determine the amount of strain introduced in the (Zn,Cd,Hg)Te-HgTe-(Zn,Cd,Hg)Te heterostructure. Magnetic field- and temperature-dependent transport measurements of Hall bar devices reveal that the change from tensile to compressive strain induces a transition from a semimetallic to a 2D-TI system for wide QWs. For thinner QWs under heavy compressive strain, a yet-unreported band gap of as high as $E_G = 55$ meV is observed.

The SLS used in this work are fabricated on GaAs:Si (001) substrates by alternating growth of CdTe in conventional MBE mode and ZnTe in atomic-layer epitaxy mode. The latter is achieved by depositing Te and Zn subsequently instead of simultaneously, which results in the self-limiting formation of a half-monolayer Zn on a monolayer Te [10]. The half filled layer of Zn atoms is completed with Cd atoms during the subsequent CdTe MBE step, resulting in

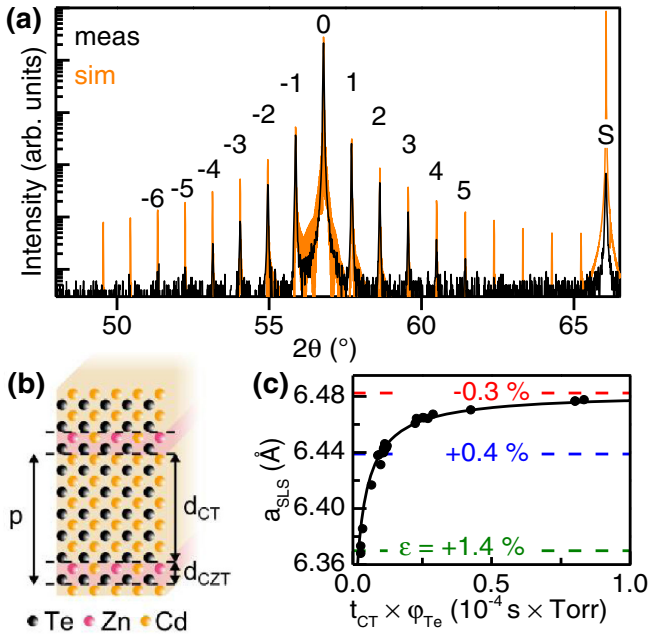

FIG. 1. (a) HRXRD $2\theta - \omega$ scan (black) and simulation (orange) of the (004) reflection of a SLS with superlattice period $p = 109$ Å. The (004) reflection from the GaAs substrate is labeled S, the SLS Bragg and higher-order peaks are labeled with 0 and $\pm 1, \pm 2, \pm 3, \ldots$, respectively. (b) Sketch of a single SLS period: one monolayer $Cd_{0.5}Zn_{0.5}Te$ alternating with an adjustable thickness d_{CT} of CdTe. (c) Dots: effective lattice constant a_{SLS} of a set of SLS with varying product of CdTe growth time and Te beam equivalent pressure $t_{CT} \times \varphi_{Te}$. The line is a fit to Eq. (1). Dashed lines indicate the lattice constants used for samples A ($\varepsilon = -0.3\%$), B ($\varepsilon = +0.4\%$), and C ($\varepsilon = +1.4\%$).

one monolayer $Cd_{0.5}Zn_{0.5}Te$ embedded in the CdTe. The layer sequence of a period of a SLS is shown in Fig. 1(b). A HRXRD $2\theta - \omega$ scan of the (004) reflection of a SLS and the simulated intensity profile are shown in Fig. 1(a). Narrow peaks and numerous satellites (labeled $\pm 1, \pm 2, \pm 3, \ldots$) indicate high crystal quality, uniform superlattice periods, and abrupt interfaces, despite the fact that growth is performed on highly lattice-mismatched GaAs substrate material. The SLS period p and, consequently, its total thickness (typically in the range of 1–3 μm) are inferred from the angular spacings of the SLS satellites. By the balancing of the forces acting within a single superlattice period and taking into account the self-limiting nature of the atomic-layer epitaxy process [11,12], one readily derives

$$a_{SLS} = a_{CT}\left[1 + \frac{f}{1 + \alpha(t_{CT} \times \varphi_{Te})}\right], \qquad (1)$$

which relates the effective lattice constant a_{SLS} of a SLS to the product of CdTe-layer growth time t_{CT} and Te beam-equivalent pressure φ_{Te}, which are both straightforwardly accessible in the experiment. In Eq. (1), f is the lattice mismatch between unstrained CdTe and $Cd_{0.5}Zn_{0.5}Te$ and a_{CT} is the lattice constant of CdTe [13]. The parameter α contains material (stiffnesses c_{ij} [14] and $Cd_{0.5}Zn_{0.5}Te$ epilayer thickness) and process-specific parameters (normalized CdTe growth speed). The effective lattice constant a_{SLS} is deduced from the angular spacing of the GaAs substrate ("S") and the zero-order Bragg reflection of the SLS ("0"). Figure 1(c) shows the obtained a_{SLS} for a set of SLS as a function of $t_{CT} \times \varphi_{Te}$. A fit of Eq. (1), with $\alpha = 3.4 \times 10^5$ s^{-1} Torr^{-1} (black line), is in good agreement with the data. Thus, a_{SLS} can be controlled over a wide range by simply adjusting $t_{CT} \times \varphi_{Te}$. This degree of freedom allows for precise control of the strain in HgTe (001) epilayers and, in turn, offers new ways to modify the band structure of bulk layers and QWs.

To demonstrate the scope of the modifications of the band structure, we have fabricated a set of three QWs, A, B, and C, with distinct strain and thickness parameters for magnetotransport measurements. Samples A and B are thick QWs with almost identical thickness ($d_{QW} = 16$ and 15 nm), and similar top and bottom barrier layers ($Cd_{0.7}Hg_{0.3}Te$ with $d_{barr} = 17$ nm, each). The virtual substrates, however, are different. Sample A is grown on a thick, relaxed CdTe (001) epilayer grown on GaAs, which gives rise to tensile strain. Samples B and C are grown on two different SLS that induce moderate and large compressive strain on the respective QWs. In sample C, a solid solution of (Zn,Cd,Hg)Te is used as barrier material ($d_{barr} = 18$ nm) (instead of the standard $Cd_{0.7}Hg_{0.3}Te$), to lower the mismatch between substrate and barriers, thus avoiding relaxation of the heterostructure. The QW thickness of sample C is $d_{QW} = 7.5$ nm. HRXRD $2\theta - \omega$ scans of the (004) diffraction profiles of all three samples are shown in Fig. 2(a). The color coding of data (red: sample A; blue: sample B; green: sample C) hold for the rest of this work. The strain in the HgTe layers, deduced from the S-0 angular separation, is $\varepsilon = -0.3\%$ for sample A, $+0.4\%$ for sample B, and $+1.4\%$ for sample C. Unlabeled reflections are caused by the (Zn,Cd,Hg)Te barriers. The barriers and the QW of all samples are fully strained, as verified by comparing the diffraction profiles with appropriate simulations (black lines). It is worth noting that the symmetric measurement geometry probes the out-of-plane response of the lattice constants to the in-plane strain. The magnitude of the response is determined by the lattice constant mismatches of SLS to barrier and SLS to QW, respectively, and the Poisson's ratios of the materials. Arrows highlight the strain-induced shift of the barrier reflection of samples A and B. Note that relaxation of the HgTe layer would be seen as a lowered shift of the top barrier reflection [15]. From its angular position, the composition of the barriers of sample C can be estimated as $Zn_{0.20}Cd_{0.56}Hg_{0.24}Te$. Remarkably, due to the large mismatch between QW and barriers, the QW of sample C is directly visible in the diffraction pattern as an isolated set of fringes [labeled "Q" in Fig. 2(a),

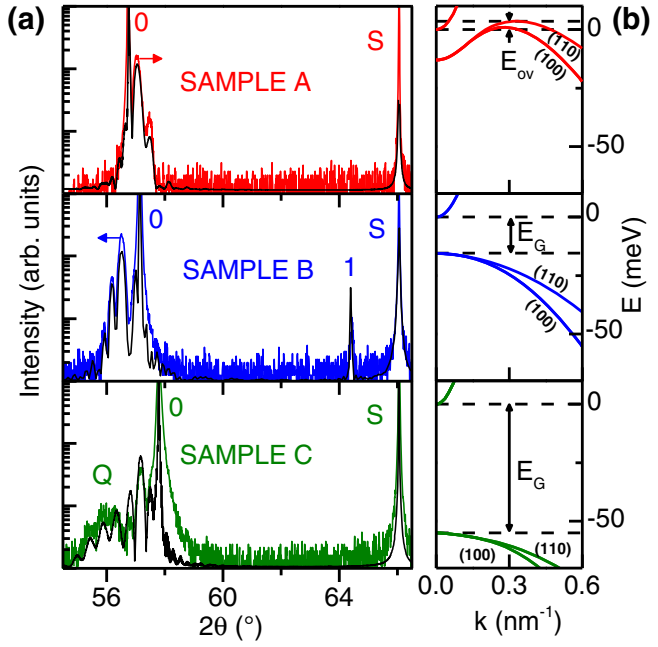

FIG. 2. (a) HRXRD $2\theta - \omega$ scan of the (004) reflection of samples A (top), B (center), and C (bottom). Reflections are labeled as in Fig. 1(a). Arrows indicate the strain-induced shift of the $Cd_{0.7}Hg_{0.3}Te$ QW barrier reflections (unlabeled). The diffracted intensity of the QW ("Q") is only visible in sample C. Black lines are simulated diffraction profiles, which are slightly offset downwards for clarity. (b) Calculated band structures of samples A, B, and C (from top to bottom). Numbers in brackets denote crystal directions. Dashed lines indicate the energetic overlap E_{ov} of VB and CB for sample A and the band gap E_G of samples B and C.

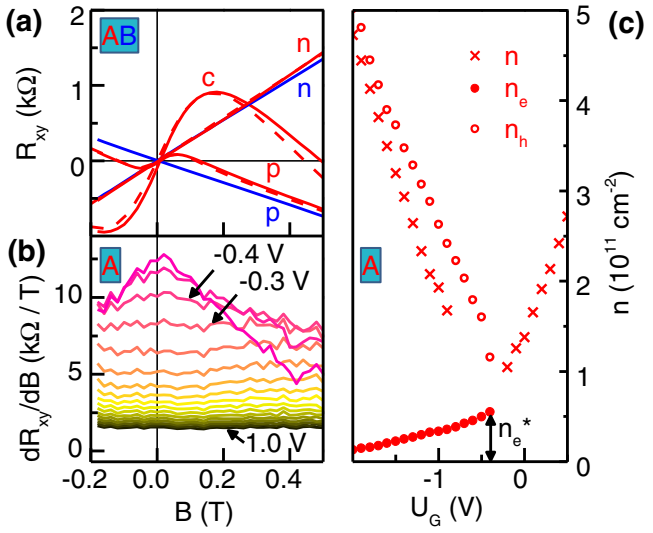

FIG. 3. (a) Hall resistance R_{xy} as a function of magnetic field for samples A (red) and B (blue) at carrier densities $n \approx +2$ and -4×10^{11} cm^{-2}, labeled n and p, respectively. Measurement at -2×10^{11} cm^{-2} (labeled c) was only possible for sample A. Dashed lines are obtained by fitting R_{xy} and R_{xx} simultaneously to a two-carrier Drude model. (b) Derivative of the Hall resistance of sample A as a function of magnetic field B for gate voltages from -0.6 V (magenta) to $+1.0$ V (black). (c) Crosses: net carrier density estimated from R_{xy} at higher fields, as a function of gate voltage U_G. Filled (empty) circles: density of electrons (holes) obtained from two-carrier fit. The double arrow highlights n-type carrier density n_e^* at the onset of two-carrier conductance.

bottom]. Figure 2(b) shows the band structures of the three QWs, calculated using an eight-band $\mathbf{k} \cdot \mathbf{p}$ model [16]. The variety in energy dispersions accessible by varying the strain and the thickness of the QW is evident. Upon comparing samples A and B, one observes that the strain in the layers primarily affects the shape of the valence band (VB) and causes a transition from a semimetal-like system with energetic overlap E_{ov} between VB and conduction band (CB) to a direct-band-gap semiconductor. From the band structure of samples B and C, one sees that the total band gap E_G increases significantly when the compressive strain is increased. Since the QW thickness of all samples is well above d_c, the band ordering is topologically nontrivial. Thus, sample A is a topological two-dimensional semimetal and samples B and C are expected to be 2D TIs.

These characteristic band dispersion properties are reflected qualitatively and quantitatively in distinct magnetotransport features. Measurements were carried out on top-gated Hall bar devices fabricated using optical lithography. In a first set of experiments (Fig. 3), we compare the behavior of samples A and B. Varying the gate voltage U_G from negative to positive values shifts the Fermi energy from the VB into the CB. This is apparent by a transition from p-conducting to n-conducting behavior that is reflected in a sign change of the Hall resistance R_{xy}. As shown in Fig. 3(a), the presence or absence of a band overlap in samples A and B results in markedly different characteristics of R_{xy} (traces with similar labels are chosen such that the carrier densities are equal within experimental resolution). When the Fermi level is deep in the CB, the Hall resistance in both samples is purely electronlike, and both traces are linear (traces labeled "n"). As the gate voltage is lowered, a pronounced curvature is observed in R_{xy} for sample A (trace "c"). This is characteristic of a system with coexisting electron- and holelike carriers of different mobilities [17], and indicates an overlap of the CB and VB [18]. In the same gate voltage regime, sample B is highly resistive, and no Hall voltage measurement is possible, implying that the Fermi energy is in the band gap. Finally, for strong negative gate voltages, an entirely linear trace is recovered for sample B, while, in contrast, two-carrier conductance persists in sample A (traces "p"). We interpret this as reflecting the effective pinning of the Fermi level at the van Hove singularity ("camel's back") in the VB density of states in sample A.

Our data allow for a more detailed analysis of the evolving electron and hole densities of sample A. As soon as the Fermi energy intersects with the VB, two-carrier

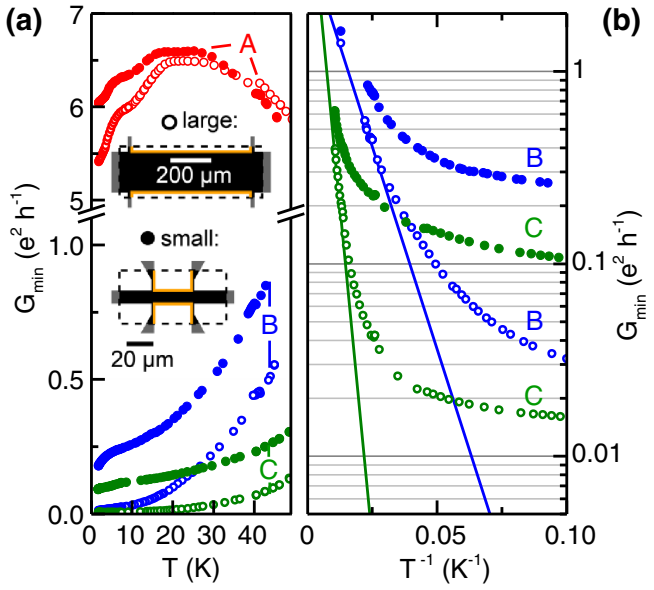

FIG. 4. (a) Minimum conductance $G_{\min}$ of samples A, B, and C as function of temperature. Empty and filled dots correspond to values measured on large and small Hall bars. Inset gives the Hall bar dimensions—note the different scales. Dashed rectangles and black areas indicate the gated regions, while the probed edge current paths are indicated by orange lines. (b) Arrhenius plot of $G_{\min}$ of samples B and C at $T > 10$ K. Lines are fit of the large Hall bar data to Eq. (2).

conductance sets in, and nonconstant dR_{xy}/dB is evidence that, at low fields, the Hall resistance is no longer described by the simple single-carrier expression $R_{xy} = B(ne)^{-1}$. Experimentally, we observe this effect for negative gate voltages larger than $U_G = -0.4$ V [Fig. 3(b)]. The onset of two-carrier conductance allows us to estimate the energy overlap E_{ov} of the VB and CB. A simultaneous fit of R_{xy} and R_{xx} to the standard two-carrier Drude model [17] yields the density of electron- and holelike carriers n_e and n_h. Fits are shown in Fig. 3(a) as dashed lines. The resulting densities for the whole measurement set are shown in Fig. 3(c), together with the net density n, extracted from R_{xy} at higher fields. The n-type carrier density at the onset of two-carrier conductance (black arrow) is $n_e^* = 5.5 \times 10^{10}$ cm^{-2}. Using $E_{\text{ov}} = E_F = n_e^* \pi \hbar^2 m_e^{-1}$, we obtain $E_{\text{ov}} = 4.5$ meV for the band overlap, which is slightly larger than the value inferred from band structure calculations [$E_{\text{ov}} = 3.5$ meV; see Fig. 2(c), top]. The electron effective mass in the CB is taken as $m_e = 0.028 m_0$ (m_0 is the mass of a free electron), in agreement with the $\mathbf{k} \cdot \mathbf{p}$ model calculations.

Finally, we discuss thermal activation studies of conductance, which allows us to discriminate between metallic sheet conductance at the charge neutrality point (sample A) and edge state conductance (samples B and C), and to estimate the magnitude of the strain-induced band gap of samples B and C. To distinguish between current flowing in the bulk of the QW and one-dimensional edge current, two Hall bars with different dimensions have been fabricated for each QW [inset in Fig. 4(a)]. Whereas the width-to-length ratio is identical ($w/l = 1/3$) for both Hall bars, the length of the gated edge changes by roughly a factor of 10 ($l_{\text{edge}} = 58$ and 620 μm). For temperatures in the range from $T = 1.8$ to 90 K, minimum values of the longitudinal conductance $G_{\min}$ are measured at gate voltages corresponding to the situation when the Fermi energy is located at the charge neutrality point for sample A and the mid-band-gap position for samples B and C. The results are plotted in Fig. 4(a). Sample A (red curves) shows a high $G_{\min}$, which changes only moderately with temperature. The observed low-temperature increase of $G_{\min}$ with temperature was reported previously in Ref. [19] and has been attributed to long-range disorder scattering [20]. We suggest that the decrease in $G_{\min}$ at higher T is due to enhanced phonon scattering. The fact that $G_{\min}$ is qualitatively similar for the large and small Hall bar indicates that the current flows in the bulk of the QW (as mentioned above, w/l, relating two-dimensional conductance to conductivity, is the same in both devices). The behavior of samples B and C is significantly different. For all temperatures, $G_{\min}$ is much smaller than in sample A, and a thermally activated increase in conductance is observed, as typical for semiconductors. A logarithmic plot of the high-temperature data ($T > 10$ K) of samples B and C versus T^{-1} is shown in Fig. 4(b). As a clear indication of edge channel transport in the low-temperature regime, we observe that $G_{\min}$ of both samples tends to saturate, and the saturation values of large and small Hall bar roughly scale with the inverse of the edge channel length (10/1). Since the edge length of both Hall bars significantly exceeds the inelastic mean-free path of the QSH edge channels [1], and thus the number of scattering events is approximately proportional to the length of the channel [21], this is an expected signature of edge channel transport. With increasing temperature, the thermally activated conductance over the whole area of the mesa becomes dominant. It is possible to extract the band gap from the conductance in the high-temperature regime. By fitting the measured $G_{\min}$ of the large Hall bar to

$$G_{\min} = G_0 \exp\left(-\frac{E_G}{2k_B T}\right), \tag{2}$$

we obtain $E_G = 17$ and 55 meV for samples B and C [solid lines in Fig. 4(b)], in good agreement with band structure calculations [$E_G = 16$ and 55 meV; see Fig. 2(b), center and bottom]. Reliable fits of Eq. (2) are only possible for the large Hall bars, since the QSH edge state conductance of the small Hall bars substantially contributes to the total conductance even at high temperatures.

In conclusion, we have presented a method to significantly increase the band gap of HgTe based 2D TIs, based on strain engineering via dedicated CdTe-$Cd_{0.5}Zn_{0.5}$Te

SLS virtual substrates. In particular, we have shown that applying compressive strain to QWs results in energy gaps as high as 55 meV. This value is the largest ever reported in inverted ($d_{QW} > 6.3$ nm) HgTe QWs, is well above $k_B T$ at room temperature, and is a necessary step towards room temperature QSH-based electronic devices. Furthermore, we have demonstrated that thick QWs can be transformed from semimetals to 2D TIs by changing their strain from tensile to compressive. Finally, we emphasize the accuracy of strain control via the SLS approach. The effective lattice constant of the SLS can be conveniently controlled by the product of CdTe-period growth time and Te beam-equivalent pressure, with both parameters straightforwardly accessible in crystal growth.

This work was supported by the German Research Foundation DFG (SPP 1666, SFB 1170), the U.S. DARPA Meso project (Grant No. N66001-11-1-4105), the European Research Council (advanced grant project 3-TOP), the Helmholtz Association (VITI), and the Elitenetzwerk Bayern.

*philipp.leubner@physik.uni-wuerzburg.de

[1] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[2] A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[3] C. Brüne, A. Roth, H. Buhmann, E. M. Hankiewicz, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Nat. Phys. **8**, 485 (2012).

[4] C. Brüne, C. X. Liu, E. G. Novik, E. M. Hankiewicz, H. Buhmann, Y. L. Chen, X. L. Qi, Z. X. Shen, S. C. Zhang, and L. W. Molenkamp, Phys. Rev. Lett. **106**, 126803 (2011).

[5] C. Brüne, C. Thienel, M. Stuiber, J. Böttcher, H. Buhmann, E. G. Novik, C.-X. Liu, E. M. Hankiewicz, and L. W. Molenkamp, Phys. Rev. X **4**, 041045 (2014).

[6] Alternatively, several μm thick buffer layers of fully relaxed CdTe on Si or GaAs can be employed.

[7] A. Pfeuffer-Jeschke, Ph.D. thesis, Universität Würzburg, 2000.

[8] R. Feldman, R. Austin, A. Dayem, and E. Westerwick, Appl. Phys. Lett. **49**, 797 (1986).

[9] R. Feldman, R. Austin, P. Fuoss, A. Dayem, E. Westerwick, S. Nakahara, T. Boone, J. Menendez, A. Pinczuk, J. Valladares *et al.*, J. Vac. Sci. Technol. B **5**, 690 (1987).

[10] Y. Takemura, M. Konagai, H. Nakanishi, and K. Takahashi, J. Cryst. Growth **117**, 144 (1992).

[11] D. Dunstan, J. Mater. Sci. **8**, 337 (1997).

[12] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.117.086403 for a brief derivation of Eq. (1).

[13] M. Schenk, I. Hähnert, L. Duong, and H.-H. Niebsch, Cryst. Res. Technol. **31**, 665 (1996).

[14] I. Andrusiv, G. Grigorovich, Y. Ilisavskii, and M. Ruvinskii, Sov. Phys. Solid State **25**, 139 (1983).

[15] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.117.086403 for additional remarks on the determination of the state of strain in the QW from HRXRD measurements and fits.

[16] E. G. Novik, A. Pfeuffer-Jeschke, T. Jungwirth, V. Latussek, C. R. Becker, G. Landwehr, H. Buhmann, and L. W. Molenkamp, Phys. Rev. B **72**, 035321 (2005).

[17] N. Ashcroft and N. Mermin, *Solid State Physics*, HRW Intl. ed. (Brooks/Cole, Belmont, MA, 1976).

[18] Z. D. Kvon, E. Olshanetsky, D. A. Kozlov, N. N. Mikhailov, and S. A. Dvoretskii, JETP Lett. **87**, 502 (2008).

[19] E. Olshanetsky, Z. D. Kvon, Y. A. Gerasimenko, V. Prudkoglyad, V. M. Pudalov, N. N. Mikhailov, and S. Dvoretsky, JETP Lett. **98**, 843 (2014).

[20] M. Knap, J. D. Sau, B. I. Halperin, and E. Demler, Phys. Rev. Lett. **113**, 186801 (2014).

[21] J. I. Väyrynen, M. Goldstein, and L. I. Glazman, Phys. Rev. Lett. **110**, 216402 (2013).